DOI:10.13656/j.cnki.gxkx.1995.01.003

广西科学 Guangxi Sciences 2 (1)：10～14 1995

非线性规划改进的广义梯度投影法*

A Modified Generalized Gradient Projection Algorithm for Nonlinear Programming

简金宝

Jian Jinbao

(广西大学数学与信息科学系 南宁市西乡塘东路10号 530004)

(Dept. of Mat. & Information Science, Guangxi University, Nanning, 530004)

摘要 讨论带非线性等式和不等式约束规划问题一种新的全局收敛的投影类算法，它是广义梯度投影法和梯度投影法以及次可行方向法的结合和改进. 算法具有广义投影阵只依赖于 ε-积极约束集，不必计算全部约束函数的梯度，其中 ε 可以随意选取；由单一的公式给出，且效益函数是可微的等优点.

关键词 非线性规划 广义梯度投影 梯度投影 全局收敛性

Abstract A new kind of global convergent projection algorithm for programming problems with nonlinear equality and inequality constraints is presented in this paper, it is a combination and modification of generalized gradient and gradient projection as well as subfeasible directions method. The algorithm possesses those main advantages, that is the generalized projection matrices only depend on ε-active constrained set and only a part of the gradient of constrained functions are computed, where ε may be chosen arbitrarily and is given by a single formula; the effective function is differentiable.

Key words Nonlinear programming, generalized gradient projection, gradient projection, global convergence

自从越民义等人建立和发展了转轴运算以来，梯度投影法得到大量深入的研究，并取得优秀的成果[1~7]. 但为保证其收敛性，每次迭代都必须通过转轴运算确定出使向量组 $\{\nabla g_j(x), j\in I\varepsilon(x)\}$ 线性无关的 ε-积极约束集 $I\varepsilon(x)$. 因而结构复杂、计算量大，影响数值的稳定性.

近年提出的广义梯度投影法[8~12]既消除了转轴运算，又保证了收敛性. 具有重要的理论意义和实用价值，不过仍存在着广义投影阵须由全部约束函数的梯度确定，因此每次迭代必须计算所有约束函数的梯度，故而投影阵的规模和计算量仍然较大及效益函数不可微等问题.

为解决上述问题，本文讨论带非线性等式与不等式约束的优化问题，引入一个只含不等式约束的可微辅助规划. 我们通过辅助规划并结合次可行方向法的思想[7,13]，建立一种更广泛的广义投影类算法，圆满地解决上述问题，计算量小、结构简单，我们将证明，算法或有限步终止于原问题的K－T点，或者产生的无穷点列的任意极限点都将是原问题的K－T点.

1 预备知识

本文考虑一般约束条件下的规划问题：

$$(P)\quad \mathrm{Min}\{f(x)\mid x\in R\}$$

其中可行集

$R=\{x\in E^n\mid g_j(x)=0, j\in L_1; g_j(x)\leqslant 0, j\in L_2\}$，$L_1\cap L_2=\Phi$，$f, g_j(j\in L_1\cup L_2)$ 为 E^n 上实值函数. 记 $L=L_1\cup L_2$.

对任意 $x\in E^n$，$\varepsilon\geqslant 0$，记

$\varphi(x)=\max\{g_j(x), j\in L\}$，$\psi(x)=\max\{0,\varphi(x)\}$，$J(x,\varepsilon)=\{j\in L_2\mid 0\leqslant\psi(x)-g_j(x)\leqslant\varepsilon\}$，$I(x,\varepsilon)=J(x,\varepsilon)\cup L_1$.

全文假设：

(a) $f, g_j(j\in L)\in C^1$；任意 $x\in E^n$，$\{\nabla g_j(x), j\in I(x,0)\}$ 线性无关.

我们引入(P)的辅助规划：

1994-08-03 收稿。

* 广西大学青年科学基金资助项目。

$(AP_c)\quad \mathrm{Min}\ F(x;c)$

$s.t.\ g_j(x) \leqslant 0, j \in L.$

其中 $F(x;c) = f(x) - c\sum_{j\in L_1} g_j(x), c > 0$ 为参数.

记(AP_c)的可行集为

$R_+ = \{x \in E^n | g_j(x) \leqslant 0, j \in L\} \supset R.$

为书写方便,记 $h(x) = \nabla f(x), h_j(x) = \nabla g_j(x)$, $j \in L; h(x;c) = \nabla F(x;c) = h(x) - c\sum_{j\in L_1} h_j(x).$

任意 $x \in E^n, \varepsilon \geqslant 0, \omega \geqslant 0$,引进如下各量

$H(x) = \mathrm{diag}[H_j(x), j \in I(x,\varepsilon)]$;

$H_j(x) = 0, j \in L_1$;

$$H_j(x) = g_j(x) - \psi(x), j \in J(x,\varepsilon) \tag{1}$$

$N(x) = (h_j(x), j \in I(x,\varepsilon))$,

$$B(x) = [N(x)^T N(x) - s(x)H(x)]^{-1} N(x)^T. \tag{2}$$

$s(x) = 1$,当 $\delta(x) = det(N(x)^T N(x)) \leqslant \omega$; $s(x) = 0$,当 $\delta(x) > \omega$. (3)

$$P(x) = E - N(x)B(x) = E - N(x)[N(x)^T N(x) - s(x)H(x)]^{-1} N(x)^T. \tag{4}$$

$u(x) = (u_j(x), j \in I(x,\varepsilon)) = -B(x)h(x)$,

$$u(x;c) = (u_j(x;c), j \in I(x,\varepsilon)) = -B(x)h(x;c) \tag{5}$$

引理 1 对任意 $x \in E^n, \varepsilon \geqslant 0, \omega \geqslant 0$,矩阵 $[N(x)^T N(x) - s(x)H(x)]$ 正定.

证明可参考文献[9,10,11].

关于原问题(P)和辅助问题(AP_c)的关系,我们有

引理 2 设参数 $c > |u_j(x)|, j \in L_1$.则 x 为原问题(P)的 $K-T$ 点当且仅当 x 为辅助问题(AP_c)的 $K-T$ 点.

证明 1° 设 x 为(P)的 $K-T$ 点.则 $x \in R \subset R_+, \psi(x) = 0$.且存在乘子 $\alpha_j, j \in I(x,\varepsilon)$,使得

$$h(x) + \sum_{j\in I(x,\varepsilon)} \alpha_j h_j(x) = 0 \tag{6}$$

$\alpha_j g_j(x) = 0, \alpha_j \geqslant 0, j \in J(x,\varepsilon)$;

$$g_j(x) = 0, j \in L_1. \tag{7}$$

于是

$$h(x;c) + \sum_{j\in J(x,\varepsilon)} \alpha_j h_j(x) + \sum_{j\in L_1} (\alpha_j + c)h_j(x) = 0. \tag{8}$$

$\alpha_j g_j(x) = 0, \alpha_j \geqslant 0, j \in J(x,\varepsilon)$;

$$(\alpha_j + c)g_j(x) = 0, j \in L_1. \tag{9}$$

由(6),(7)有 $h(x) + N(x)\alpha = 0$, $s(x)H(x)\alpha = 0$.故

$N(x)^T h(x) + N(x)^T N(x)\alpha - s(x)H(x)\alpha = 0$,

$\alpha = -[N(x)^T N(x) - s(x)H(x)]^{-1} N(x)^T h(x) = u(x)$.

由 $c > |u_j(x)|, j \in L_1$.知 $\alpha_j + c = u_j + c > 0$, $j \in L_1$.于是由(8),(9)可知 x 为(AP_c)的 $K-T$ 点.

2° 设 x 为(AP_c)的 $K-T$ 点,则存在 $\beta_j, j \in I(x,\varepsilon)$,使得

$$h(x;c) + \sum_{j\in J(x,\varepsilon)} \beta_j h_j(x) + \sum_{j\in L_1} \beta_j h_j(x) = 0. \tag{10}$$

$\beta_j g_j(x) = 0, \beta_j \geqslant 0, j \in I(x,\varepsilon)$;

$$\psi(x) = 0, x \in R_+. \tag{11}$$

故

$$h(x) + \sum_{j\in J(x,\varepsilon)} \beta_j h_j(x) + \sum_{j\in L_1} (\beta_j - c)h_j(x) = 0. \tag{12}$$

$$\beta_j g_j(x) = 0, \beta_j \geqslant 0, j \in I(x,\varepsilon).\ x \in R_+. \tag{13}$$

记

$\gamma_j = \beta_j, j \in J(x,\varepsilon)$;

$\gamma_j = \beta_j - c, j \in L_1$.

$\gamma = (\gamma_j, j \in I(x,\varepsilon))$.

则 $h(x) + N(x)\gamma = 0$.

又由(13)有 $s(x)H(x)\gamma = 0$.

故 $N(x)^T h(x) + N(x)^T N(x)\gamma - s(x)H(x)\gamma = 0, \gamma = u(x)$.因此 $\beta_j = \gamma_j + c = u_j(x) + c > 0, j \in L_1$.由(13)有 $g_j(x) = 0, j \in L_1$.从而 $x \in R$,结合(12),(13)可知 x 为(P)的 $K-T$ 点.

证毕.

为简炼以后的讨论,我们先给出以下关系式,这些式子均可利用简单的凑项技巧和各量的定义给予证明,也可参考文献[9～12].

$$N(x)^T P(x) = -s(x)H(x)[N(x)^T N(x) - s(x)H(x)]^{-1} N(x)^T \tag{14}$$

$$N(x)^T B(x)^T = E + s(x)H(x)[N(x)^T N(x) - s(x)H(x)]^{-1} \tag{15}$$

$P(x)h(x;c) = P(x)h(x)$,

$$h(x;c)^T P(x)h(x) = h(x)^T P(x)h(x) \geqslant \| P(x)h(x) \|^2 \tag{16}$$

$u_j(x;c) = u_j(x), j \in J(x,\varepsilon)$;

$$u_j(x;c) = u_j(x) + c, j \in L_1. \tag{17}$$

引理 3 对于 $x \in E^n$, x 为(AP_c)的 $K-T$ 点,当且仅当

$P(x)h(x) = 0$,

$u_j(x;c)g_j(x) = 0, u_j(x;c) \geqslant 0, j \in I(x,\varepsilon)$;

$$\psi(x) = 0. \tag{18}$$

证明　如(18)成立,利用(16)有

$$P(x)h(x;c) = h(x;c) + \sum_{j\in I(x,\varepsilon)} u_j(x;c)h_j(x) = 0 \quad (19)$$

$$u_j(x;c)g_j(x) = 0,$$

$$u_j(x;c) \geq 0, j \in I(x,\varepsilon); x \in R_+. \quad (20)$$

由此可知 x 为 (AP_c) 的 $K-T$ 点.

反之,如 x 为 (AP_c) 的 $K-T$ 点,由引理2之2°的证明和(17)知,此时(19),(20)成立,从而(18)成立.

证毕.

现在我们可以构造搜索方向 $d(x)$,记

$$\alpha(x) = \sum_{j\in J(x,\varepsilon)} min\{0,u_j(x),u_j(x)(g_j(x)-\psi(x))\} + \sum_{j\in L_1}(u_j(x)+c)(g_j(x)-\psi(x)). \quad (21)$$

$$\beta(x) = \sum_{j\in I(x,\varepsilon)} |u_j(x)| + c\|L\|, \tau(x) = -\|P(x)h(x)\|^2 + \alpha(x) - \psi(x). \quad (22)$$

$$\rho(x) = -\tau(x)/(1+\beta(x)),$$

$$d(x) = -P(x)h(x) + B(x)^T\sigma(x). \quad (23)$$

$$\sigma_j(x) = \begin{cases} -1-\rho(x), \text{如 } j \in J(x,\varepsilon), u_j(x) < 0, \\ \psi(x) - g_j(x) - \rho(x), \text{如 } j \in J(x,\varepsilon), \\ \qquad u_j(x) \geq 0; \text{或 } j \in L_1. \end{cases} \quad (24)$$

引理4　设 $c > |u_j(x)|; j \in L_1$. 则 x 为 (P) 的 $K-T$ 点,当且仅当 $\rho(x) = 0$.

定理1　$h(x;c)^Td(x) \leq \psi(x) - \rho(x)$, $h_j(x)^Td(x) \leq -\rho(x)$,对 $j \in L$,且 $g_j(x) = \psi(x)$.

证明　$h(x;c)^Td(x) = -h(x;c)^TP(x)h(x) - \sum_{j\in I(x,\varepsilon)} u_j(x;c)\sigma_j(x)$,由(16,17,24)有

$$h(x;c)^Td(x) - \|P(x)h(x)\|^2 + \rho(x)\sum_{j\in J(x,\varepsilon)} u_j(x) + \sum_{j\in J(x,\varepsilon),u_j(x)\geq 0} u_j(x)[g_j(x)-\psi(x)]$$

$$+ \sum_{j\in J(x,\varepsilon),u_j(x)<0} u_j(x) + \sum_{j\in L_1}(u_j(x)+c)(g_j(x) - \psi(x)) + \rho\sum_{j\in L_1} u_j(x) + \rho c\|L\|$$

$$\leq -\|P(x)h(x)\|^2 + \rho\sum_{j\in I(x,\varepsilon)} |u_j(x)| + \rho c\|L\| + \alpha(x),$$

$$h(x;c)^Td(x) \leq -\|P(x)h(x)\|^2 + P(x)\beta(x) + \alpha(x) = -\rho(x) + \psi(x).$$

对于 $j \in L, g_j(x) = \psi(x)$,有 $j \in I(x,\varepsilon)$,且 $H_j(x) = 0$. 由(14,15,24)有

$$h_j(x)^Td(x) = \sigma_j(x) = -1-\rho(x), \text{或} -\rho(x).$$

故 $h_j(x)^Td(x) \leq -\rho(x)$.

关于假设(a),它比一般非退化假设,即 $\{h_j(x), j \in L_1 \cup \{j: g_j(x) = \max\{0,g_i(x), i \in L_2\}\}$,线性无关,要弱. 事实上,要么 $J(x,0) = \Phi$,要么 $J(x,0) = \{j \in L_2: g_j(x) = \max\{0,g_i(x), i \in L_2\}\}$.

2　算法及性质

在以上预备知识基础上,我们给出问题(P)改进的广义投影算法的迭代步骤.

算法步骤:

步0　任取 $x^1 \in E^n, \gamma > 0, c_0 > 0, k := 1$.

步1　对 x^k,选取 $\varepsilon_k \geq 0, \omega_k \geq 0$,计算 $\psi_k = \psi(x^k), J_k = J(x^k,\varepsilon_k), I_k = J_k \cup L_1, N_k = N(x^k), H_k = H(x^k), \delta_k = \delta(x^k), s_k = s(x^k)$. 按(2),(4),(5)计算 $B_k = B(x^k), P_k = P(x^k), u^k = u(x^k)$.

步2　调整参数 c:如 $L_1 = \Phi$,则 $c_k = c_0$. 否则计算 $q_k = \max\{|u_j^k|, j \in L_1\} + c_0$,

如 $q_k > c_{k-1}$,令 $c_k = \max\{q_k, c_{k-1} + \gamma\}$;

如 $q_k \leq c_{k-1}$,令 $c_k = c_{k-1}$.

步3　按(21)~(24)计算 $\alpha_k = \alpha(x^k), \beta_k = \beta(x^k), \tau_k = \tau(x^k), \rho_k = \rho(x^k)$, $\sigma_k = (\sigma_j^k, j \in I_k) = \sigma(x^k), d^k = d(x^k)$. 如 $\rho_k = 0$,则 x^k 为原问题(P)的 $K-T$ 点,停.

步4　求 $\{1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots\}$ 中满足(25),(26)的最大值 λ_k.

$$F(x^k + \lambda d^k; c_k) \leq F(x^k; c_k) + \frac{1}{2}\lambda(h(x^k;c_k)^Td^k + \psi_k) \quad (25)$$

$$g_j(x^k + \lambda d^k) - \psi_k \leq -\frac{1}{2}\lambda\rho_k, \forall j \in L = L_1 \cup L_2. \quad (26)$$

步5　令 $x^{k+1} = x^k + \lambda_k d^k, k := k+1$,返回步1.

由于 $c_k \geq q_k > |u_j^k|, j \in L_1$,由引理4知,上述算法终止时,$x^k$ 为(P)的 $K-T$ 点. 为说明线搜索步4可在有限次终止,我们有

引理5　对于充分小的正数 λ,(25),(26)式成立.

引理5可由定理1及(25),(26)获证,省略.

由定理1和(25),(26),我们易知:当 $x^k \in R_+$,即 $\psi_k = 0$ 时,d^k 为 (APc_k) 的可行下降方向,且 $F(x^{k+1}; c_k) < F(x^k; c_k), x^{k+1} \in R_+$. 从而一旦有 $x^t \in R_+$,则任意 $k \geq t, x^k \in R_+$.

从算法有效的角度说,参数 $\varepsilon_k \geq 0, \omega_k \geq 0$ 可任意选取. 为便于其实际应用,我们给出几种选取法,从而产生一些有趣的方法. 其中 $\varepsilon_0 > 0$.

取法 I　$\omega_k = \varepsilon_0 + \delta_k > \delta_k, \varepsilon_k = \varepsilon_0 + \max\{\psi_k -$

$g_j(x^k), j \in L\} \geqslant \varepsilon_0$.

此时 $s_k \equiv 1, J_k = J(x^k, \varepsilon_k) \equiv L_2, I_k \equiv L_2 \cup L_1 = L$. P_k 为通常的广义投影阵[9-12].

取法 Ⅰ　$\omega_k = \varepsilon_0 + \delta_k > \delta_k, \varepsilon_k \equiv \varepsilon_0$.

此时 $s_k \equiv 1, P_k$ 是仅由 ε_0— 积极约束集确定的广义投影阵. 特别当 ε_0 非常小时, $J(x_k, \varepsilon_k)$ 和 $J(x^k, 0)$ 十分接近.

取法 Ⅱ　$\varepsilon_k = \omega_k$ 用通常转轴运算产生. 使得 $\delta_k > \omega_k$.

此时 $s_k \equiv 0, P_k$ 为通常投影阵, 算法成为一般梯度投影法.

取法 Ⅳ. $\varepsilon_k = \varepsilon_0, \omega_k = \omega_0 > 0$. 此时算法是梯度投影法和广义梯度投影法的结合.

取法 Ⅴ. $\varepsilon_k = \omega_k = 0$, 由 (a) 知 $s_k \equiv 0, P_k$ 只由主动约束集产生.

由上可见, 本文处理方法具有普遍意义.

3　算法收敛性论证

当算法有限步终止于 x^k 时, 由引理 4 知 x^k 为 (P) 的 $K-T$ 点. 在以下的收敛性讨论中, 不妨设算法产生无穷点列 $\{x^k\}$, 我们将证明 $\{x^k\}$ 的任意极限点 x^* 必是原问题 (P) 的 $K-T$ 点. 为保证收敛性, 需对 ε_k, ω_k 作如下假设:

(F): 存在 $\varepsilon > 0, \omega > 0$, 使得 k 充分大时 $\varepsilon_k \geqslant \varepsilon$, $\omega_k \geqslant \omega$.

易见 ε_k, ω_k 的取法 Ⅰ—Ⅳ 均满足 (F).

由于 $J_k \subset L_2$, 不妨设存在子列 K 使得 $x^k \to x^*$, $k \in$ K, $J_k \equiv J, k \in$ K. $I_k \equiv J \cup L_1 = I$. 又 s_k 只取 0, 1, 可不妨假设 $s_k \equiv 0 = s, k \in$ K; 或 $s_k \equiv 1 = s, k \in$ K.

令 $s(x^*) = s$. 因当 $s = 0$ 时,

$$\det(N(x^*)^T N(x^*)) = \operatorname*{limit}_{k \in K} \delta_k \geqslant \operatorname*{limit}_{k \in K} \omega_k \geqslant \omega > 0.$$

由此易见 $N(x^*)^T N(x^*) - s(x^*) H(x^*)$ 可逆, 其中 $N(x^*) = (h_j(x^*), j \in I)$.

引理 6　存在正整数 t, 使得 $c_k \equiv c_t = c$, $\forall\, k \geqslant t$. (证明可参考文献 7, 10).

对于 x^*, 我们用 (21), (22), (23) 计算 $\alpha(x^*)$, $\beta(x^*), \tau(x^*), \rho(x^*)$. 由于 $\{\sigma(x^k)\}$ 是有界的 ($k \in$ K), 可不妨设 $\sigma_k = \sigma(x^k) \to \sigma^* = (\sigma_j^*, j \in I), k \in$ K. 令

$d^* = -P(x^*)h(x^*) + B(x^*)^T \sigma^*$. 则 $\{\rho_k, d^k\} \to \{\rho(x^*), d^*\}, k \in$ K.

引理 7　如 x^* 为非 $K-T$ 点, 则 $\rho(x^*) > 0$, $h(x^*; c)^T d^* - \psi(x^*) \leqslant -\rho(x^*) < 0$.

证　由于 $c_k = c \geqslant q_k \geqslant |u_j^k| + c_0$, 故 $c \geqslant |u_j^*| + c_0 > |u_j^*|, j \in L_1$.
利用引理 4 知 $\rho(x^*) > 0$. 又由定理 1 有

$$\begin{aligned} h(x^*; c)^T d^* &= \operatorname*{limit}_{k \in K} h(x^k; c)^T d^k \\ &\leqslant \operatorname*{limit}_{k \in K} (\psi(x^k) - \rho(x^k)) \\ &= \psi(x^*) - \rho(x^*). \end{aligned}$$

引理 7 成立.

引理 8　如 x^* 不是 $K-T$ 点, 则 $\lambda_0 = \inf\{\lambda_k, k \in$ K$\} > 0$.

证　a) 对 $g_j(x^*) < \psi(x^*), j \in L$ 的 j: 有 $g_j(x^k) - \psi(x^k) \to g_j(x^*) - \psi(x^*) = -2\xi < -\xi < 0, k \in$ K. 所以当 $k \in$ K 充分大时, $g_j(x^k) - \psi(x^k) < -\xi$. 因此.

$$\begin{aligned} & g_j(x^k + \lambda d^k) - \psi(x^k) + \frac{1}{2}\lambda\rho_k \\ &= \lambda h_j(x^k)^T d^k + g_j(x^k) - \psi(x^k) + \frac{1}{2}\lambda\rho_k + o(\lambda) \\ &= g_j(x^k) - \psi(x^k) + O(\lambda) < -\xi + O(\lambda). \end{aligned}$$

于是对充分小的 λ 及充分大的 $k \in$ K, $j \in L$: $g_j(x^*) < \psi(x^*)$, (26) 式成立时.

b)　对 $j \in L, g_j(x^*) = \psi(x^*)$: 由于 $j \in L \backslash (J_K \cup L_1)$ 时, $\psi(x^k) - g_j(x^k) > \varepsilon_k \geqslant \varepsilon > 0$. 故 $\psi(x^*) - g_j(x^*) \geqslant \varepsilon > 0$. 所以当 $g_j(x^*) = \psi(x^*)$ 时, $j \in J_k \cup L_1 = I_k = I$. 由 $H_j(x^*) = 0, j \in L_1, H_j(x^*) = g_j(x^*) - \psi(x^*), j \in J = J_k$. 有 $H_j(x^*) = 0$.

利用 (14), (15), (24), 易证

$h_j(x^*)^T d^* = \sigma_j^*$.

由 (25) 及 $\sigma_j^k \to \sigma_j^*$, 易见 $\sigma_j^* = -1 - \rho(x^*)$ 或 $\sigma_j^* = -\rho(x^*) + \psi(x^*) - g_j(x^*) = -\rho(x^*)$. 所以 $h_j(x^*)^T d^* = \sigma_j^* \leqslant -\rho(x^*) < 0$. 于是

$$\begin{aligned} & h_j(x^k)^T d^k \to h_j(x^*)^T d^* \\ & \leqslant -\rho(x^*) < -\frac{3}{4}\rho(x^*). \end{aligned}$$

故对充分大的 $k \in$ K, 有 $h_j(x^k)^T d^k \leqslant -\dfrac{3}{4}\rho(x^*)$.

$$\begin{aligned} & g_j(x^k + \lambda d^k) - \psi(x^k) + \frac{1}{2}\lambda\rho(x^k) \\ & \leqslant g_j(x^k + \lambda d^k) - g_j(x^k) + \frac{1}{2}\lambda\rho(x^k) \\ & = \lambda h_j(x^k)^T d^k + \frac{1}{2}\lambda\rho(x^k) + o(\lambda) \\ & \leqslant -\frac{3}{4}\lambda\rho(x^*) + \frac{1}{2}\cdot\frac{4}{3}\lambda\rho(x^*) + o(\lambda) \\ & = -\frac{1}{12}\lambda\rho(x^*) + o(\lambda). \end{aligned}$$

故对 $\lambda > 0$ 充分小时,

$$g_j(x^k + \lambda d^k) - \psi(x^k) \leqslant -\frac{1}{12}\lambda\rho(x^k).$$

即对 $g_j(x^*)=\psi(x^*)$ 的 j，当 $\lambda>0$ 充分小，$k\in K$ 充分大时，(26) 式成立.

C) 由于 $h(x^k;c)^Td^k-\psi(x^k)\to h(x^*;c)^Td^*-\psi(x^*)\leqslant-\rho(x^*)<0$，

故当 $k\in K$ 充分大时，

$h(x^k;c)^Td^k-\psi(x^k)<-\frac{1}{2}\rho(x^*)$. 于是

$$F(x^k+\lambda d^k;c)-F(x^k;c)-\frac{1}{2}\lambda h(x^k;c)^Td^k-\frac{1}{2}\lambda\psi(x^k)=\frac{1}{2}\lambda(h(x^k;c)^Td^k-\psi(x^k))+0(\lambda)$$
$$\leqslant-\frac{1}{2}\lambda\rho(x^*)+o(\lambda).$$

从而 $k\in K$ 充分大，$\lambda>0$ 充分小时，

$$F(x^k+\lambda d^k;c)\leqslant F(x^k;c)+\frac{1}{2}\lambda(h(x^k;c)^Td^k+\psi(x^k)),$$

即(25) 式成立.

综合上述 a),b),c) 及步 4 可知引理 8 成立.

定理 2 在假设(a) 及(F) 下，$\{x^k\}$ 的任意极限点 x^* 必是原问题(P) 的 $K-T$ 点.

证明 假若定理不成立，则由引理 8 及(25)，(26) 易见 $\{x^k\}$ 具有如下两种性质之一：

性质 Ⅰ. 存在 $x^t\in R_+$. 从而 $x^k\in R_+$，$\psi(x^k)=0$，$\forall k\geqslant t$，且

$$F(x^{k+1};c)\leqslant F(x^k;c)+\frac{1}{2}\lambda_kh(x^k;c)^Td^k\leqslant F(x^k;c)-\frac{1}{2}\lambda_k\rho(x^k)\leqslant F(x^k;c)-\frac{1}{2}\lambda_0\rho(x^*).$$

于是 $\{F(x^k;c)\}$ 单调下降，由 $F(x^k;c)\to F(x^*;c)$，$k\in K$，有 $\lim\limits_{k\to\infty}F(x^k;c)=F(x^*;c)$

在上面不等式中令 $k\in K$，$k\to\infty$，有 $-\frac{1}{2}\lambda_0\rho(x^*)\geqslant0$，这与引理 7,8 矛盾

性质 Ⅱ. $\forall k=1,2,3,\dots$，$x^k\overline{\in}R_+$. 从而 $\varphi(x^k)=\psi(x^k)>0$.

由(26) 式知

$$g_j(x^{k+1})-\psi(x^k)=g_j(x^{k+1})-\phi(x^k)\leqslant-\frac{1}{2}\lambda_k\rho(x^k)\leqslant-\frac{1}{2}\lambda_0\rho(x^k).$$

$$\varphi(x^{k+1})=\max\{g_j(x^{k+1}),j\in L\}\leqslant\varphi(x^k)-\frac{1}{2}\lambda_k\rho(x^k).$$

故 $\{\varphi(x^k)\}$ 为单调下降的. 又 $\varphi(x^k)\to\varphi(x^*)$，$k\in K$，所以 $\varphi(x^k)\to\varphi(x^*)$，$k\to\infty$.

在 $\varphi(x^{k+1})\leqslant\varphi(x^k)-\frac{1}{2}\lambda_k\rho(x^k)\leqslant\varphi(x^k)-\frac{1}{2}\lambda_0\rho(x^k)$ 中令 $k\to\infty$，$k\in K$ 有 $-\frac{1}{2}\lambda_0\rho(x^*)\geqslant0$，这也矛盾于引理 7,8.

由此可知定理 2 成立.

证毕.

最后，关于线搜索(25)，(26)，我们作以下注明：

注 1 当存在 $x^t\in R_+$ 时，由(25)，(26) 产生的点列 $\{x^k\}(k\geqslant t)$ 为严格内点. 故在实际近似计算中，为尽快停机，可根据精确度要求，当 $\rho(x^k)$ 足够小时可终止算法.

注 2 线搜索(25,26) 可用(25,28)，(26,27)，(27,28) 之一代替，同样具有收敛性

$$(1-r_k)[F(x^k+\lambda d^k;c)-F(x^k;c)-\frac{1}{2}\lambda h(x^k;c)^Td^k]\leqslant0 \tag{27}$$

$$g_j(x^k+\lambda d^k)-\psi(x^k)\leqslant-\frac{1}{2}r_k\lambda\rho(x^k),\forall j\in L. \tag{28}$$

其中 $r_k=1$，当 $\varphi(x^k)>0$；$r_k=0$，当 $\varphi(x^k)=0$.

参考文献

1 Yue Miyi and Han Jiye, A new reduced gradient method, Scientia Sinica, 1979, 22: 10: 1099～2113.

2 Du Dingzhu, Sun Jie and Song Tiantan, Simplified finite pivoting processes in reduced gradiend algorithm, ACTA Mathematical Applicate Sinica, 1984, 7: 2.

3 赖炎连．非线性约束凸规划的一个解法及其收敛性．应用数学学报，1980，4，322～331.

4 赖炎连．系统科学与数学，1990，10，2：181～188.

5 堵丁柱．非线性约束条件下的梯度投影方法，应用数学学报，1985，8，1：7～16.

6 薛声家．解非线性约束拟凸规划的一个梯度投影法，数学研究与评论，2（1984），87～92.

7 简金宝．非线性约束规划问题的一个全局收敛的次可行方向法．曲阜师范大学学报，1992，18：4：55～61.

8 Jose Herskovits, Math. Prog., 1986, 36, 1: 19～38.

9 高自友，贺国平．约束优化问题的一个广义梯度投影法，科学通报，1991，36：19，1443～1447.

10 赖炎连，高自友，贺国平．非线性最优化的广义梯度投影法．中国科学，9（A），1992：916～924.

11 高自友，于 伟．任意初始点下的广义梯度投影方法．科学通报，1992，37：20：1833～1836.

12 赖炎连，马红缨．优化问题的一个广义梯度投影法．见运筹与决策（Vol. 1），ORSC'92，成都：成都科技大学出版社，1992：453～460.

13 简金宝．次可行和强次可行方向法的研究．见：最优化理论与应用．西安电子科技大学出版社，1994 年 10 月，113～118.

（责任编辑：莫鼎新、何启彬）